

# PATENT ABSTRACTS OF JAPAN

(11) Publication number : 2002-258886  
 (43) Date of publication of application : 11.09.2002

(51) Int.Cl.

G10L 13/08  
 A63H 11/00  
 // B25J 5/00  
 B25J 13/00

(21) Application number : 2001-058843

(22) Date of filing : 02.03.2001

(71) Applicant : SONY CORP

(72) Inventor : YAMAZAKI NOBUHIDE

AKAHA MAKOTO

YAMADA KEIICHI

KOBAYASHI KENICHIRO

NITSUTA TOMOAKI

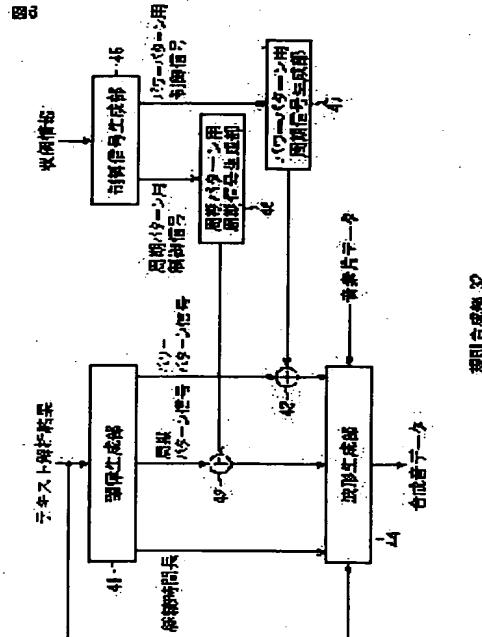
KOBAYASHI ERIKA

## (54) DEVICE AND METHOD FOR COMBINING VOICES, PROGRAM AND RECORDING MEDIUM

### (57) Abstract:

**PROBLEM TO BE SOLVED:** To easily generate a combined voice corresponding to a state, such as feelings by easily controlling the rhythm of a combined voice.

**SOLUTION:** A periodic signal for periodic pattern generating part 46 and a periodic signal for power pattern generating part 47 generate a periodic signal, that periodically changes rhythm information, and computing devices 42 and 43 correct the rhythm information, by adding the periodical signal and the rhythm information. A waveform-generating part 44 performs voice combination, by using the corrected rhythm information obtained from the result.



### LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision  
of rejection]

[Date of requesting appeal against examiner's  
decision of rejection]

[Date of extinction of right]

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号  
特開2002-258886  
(P2002-258886A)

(43)公開日 平成14年9月11日(2002.9.11)

(51)Int.Cl.  
G 1 0 L 13/08  
A 6 3 H 11/00  
// B 2 5 J 5/00  
13/00

識別記号

F I  
A 6 3 H 11/00  
B 2 5 J 5/00  
13/00  
G 1 0 L 3/00

テ-マコト\*(参考)  
Z 2 C 1 S 0  
C 3 C 0 0 7  
Z 5 D 0 4 5  
H

審査請求 未請求 請求項の数8 OL (全14頁)

(21)出願番号 特願2001-58843(P2001-58843)

(22)出願日 平成13年3月2日(2001.3.2)

(71)出願人 000002185  
ソニー株式会社  
東京都品川区北品川6丁目7番35号  
(72)発明者 山崎 信英  
東京都品川区北品川6丁目7番35号 ソニ  
ー株式会社内  
(72)発明者 赤羽 誠  
東京都品川区北品川6丁目7番35号 ソニ  
ー株式会社内  
(74)代理人 100082131  
弁理士 稲本 義雄

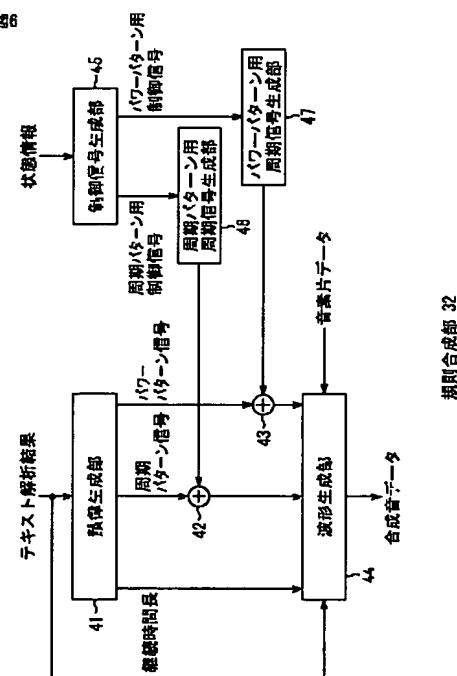
最終頁に続く

(54)【発明の名称】 音声合成装置および音声合成方法、並びにプログラムおよび記録媒体

(57)【要約】

【課題】 合成音の韻律を容易に制御し、これにより、感情等の状態に応じた合成音を、容易に生成する。

【解決手段】 周期パターン用周期信号生成部46とパワーパターン用周期信号生成部47において、韻律情報を周期的に変化させる周期信号が生成され、演算器42と43において、周期信号と韻律情報を加算することにより、韻律情報が補正される。そして、波形生成部44では、その結果得られる補正韻律情報を用いて音声合成が行われる。



【特許請求の範囲】

- 【請求項1】 少なくとも韻律情報を用いて音声合成を行う音声合成装置であって、前記韻律情報を周期的に変化させる周期信号を生成する周期信号生成手段と、前記周期信号と韻律情報を加算することにより、前記韻律情報を補正し、補正韻律情報を出力する補正手段と、前記補正韻律情報を用いて音声合成を行う音声合成手段とを備えることを特徴とする音声合成装置。
- 【請求項2】 所定の制御信号を生成して出力する制御信号生成手段をさらに備え、前記周期信号生成手段は、前記制御信号に基づいて、周波または振幅を制御した前記周期信号を生成することを特徴とする請求項1に記載の音声合成装置。
- 【請求項3】 前記制御信号生成手段は、外部から供給される、感情の状態を表す感情モデルに基づいて、前記制御信号を生成することを特徴とする請求項2に記載の音声合成装置。
- 【請求項4】 前記韻律情報は、ピッチ周期のパターンを表す周期パターン信号、または合成音のパワーのパターンを表すパワーパターン信号であることを特徴とする請求項1に記載の音声合成装置。
- 【請求項5】 前記周期信号生成手段は、音節の長さ以上の周期の前記周期信号を生成することを特徴とする請求項1に記載の音声合成装置。
- 【請求項6】 少なくとも韻律情報を用いて音声合成を行う音声合成方法であって、前記韻律情報を周期的に変化させる周期信号を生成する周期信号生成ステップと、前記周期信号と韻律情報を加算することにより、前記韻律情報を補正し、補正韻律情報を出力する補正ステップと、前記補正韻律情報を用いて音声合成を行う音声合成ステップとを備えることを特徴とする音声合成方法。
- 【請求項7】 少なくとも韻律情報を用いて音声合成を行う音声合成処理を、コンピュータに行わせるプログラムであって、前記韻律情報を周期的に変化させる周期信号を生成する周期信号生成ステップと、前記周期信号と韻律情報を加算することにより、前記韻律情報を補正し、補正韻律情報を出力する補正ステップと、前記補正韻律情報を用いて音声合成を行う音声合成ステップとを備えることを特徴とするプログラム。
- 【請求項8】 少なくとも韻律情報を用いて音声合成を行う音声合成処理を、コンピュータに行わせるプログラムが記録されている記録媒体であって、前記韻律情報を周期的に変化させる周期信号を生成する周期信号生成ステップと、

前記周期信号と韻律情報を加算することにより、前記韻律情報を補正し、補正韻律情報を出力する補正ステップと、

前記補正韻律情報を用いて音声合成を行う音声合成ステップとを備えるプログラムが記録されていることを特徴とする記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、音声合成装置および音声合成方法、並びにプログラムおよび記録媒体に関し、特に、合成音の韻律を容易に制御し、これにより、例えば、感情等の状態に応じた合成音を、容易に生成することができるようになる音声合成装置および音声合成方法、並びにプログラムおよび記録媒体に関する。

【0002】

【従来の技術】従来の音声合成装置においては、韻律モデルに基づいて、合成音の韻律情報が生成され、さらに、その韻律情報にしたがって、合成音のピッチの周波パターンや、合成音のパワーパターン等の韻律を制御した合成音が生成される。

【0003】

【発明が解決しようとする課題】ところで、最近、例えば、ペット型のペットロボット等として、音声合成装置を搭載し、ユーザーに話しかけるものが提案されている。

【0004】さらに、ペットロボットとしては、感情の状態を表す感情モデルを取り入れ、その感情モデルが表す感情の状態に応じて、ユーザーの命令に従ったり、従わなかったりするものも提案されている。

【0005】従って、感情モデルに応じて、合成音の韻律を変化させることができれば、感情に応じた合成音が输出されることから、ペットロボットのエンタテインメント性を向上させることができると考えられる。

【0006】しかしながら、従来の音声合成装置においては、合成音の韻律は、韻律モデルのパラメータを制御することにより変化させていたため、複雑な制御が必要であった。即ち、感情を表現する合成音を生成する場合には、感情モデルに応じて、韻律モデルの各種のパラメータを制御する必要があり、面倒であった。

【0007】本発明は、このような状況に鑑みてなされたものであり、合成音の韻律を容易に制御し、これにより、感情等の状態に応じた合成音を、容易に生成することができるようになるものである。

【0008】

【課題を解決するための手段】本発明の音声合成装置は、韻律情報を周期的に変化させる周期信号を生成する周期信号生成手段と、周期信号と韻律情報を加算することにより、韻律情報を補正し、補正韻律情報を出力する補正手段と、補正韻律情報を用いて音声合成を行う音声合成手段とを備えることを特徴とする。

【0009】本発明の音声合成方法は、韻律情報を周期

的に変化させる周期信号を生成する周期信号生成ステップと、周期信号と韻律情報を加算することにより、韻律情報を補正し、補正韻律情報を出力する補正ステップと、補正韻律情報を用いて音声合成を行う音声合成ステップとを備えることを特徴とする。

【0010】本発明のプログラムは、韻律情報を周期的に変化させる周期信号を生成する周期信号生成ステップと、周期信号と韻律情報を加算することにより、韻律情報を補正し、補正韻律情報を出力する補正ステップと、補正韻律情報を用いて音声合成を行う音声合成ステップとを備えることを特徴とする。

【0011】本発明の記録媒体は、韻律情報を周期的に変化させる周期信号を生成する周期信号生成ステップと、周期信号と韻律情報を加算することにより、韻律情報を補正し、補正韻律情報を出力する補正ステップと、補正韻律情報を用いて音声合成を行う音声合成ステップとを備えるプログラムが記録されていることを特徴とする。

【0012】本発明の音声合成装置および音声合成方法、並びにプログラムにおいては、韻律情報を周期的に変化させる周期信号が生成され、周期信号と韻律情報を加算することにより、韻律情報が補正される。そして、その結果得られる補正韻律情報を用いて音声合成が行われる。

### 【0013】

【発明の実施の形態】図1は、本発明を適用したロボットの一実施の形態の外観構成例を示しており、図2は、その電気的構成例を示している。

【0014】本実施の形態では、ロボットは、例えば、犬等の四つ足の動物の形状のものとなっており、胴体部ユニット2の前後左右に、それぞれ脚部ユニット3A、3B、3C、3Dが連結されるとともに、胴体部ユニット2の前端部と後端部に、それぞれ頭部ユニット4と尻尾部ユニット5が連結されることにより構成されている。

【0015】尻尾部ユニット5は、胴体部ユニット2の上面に設けられたベース部5Bから、2自由度をもって湾曲または摆動自在に引き出されている。

【0016】胴体部ユニット2には、ロボット全体の制御を行うコントローラ10、ロボットの動力源となるバッテリ11、並びにバッテリセンサ12および熱センサ13からなる内部センサ部14などが収納されている。

【0017】頭部ユニット4には、「耳」に相当するマイク（マイクロフォン）15、「目」に相当するCCD（Charge Coupled Device）カメラ16、触覚に相当するタッチセンサ17、「口」に相当するスピーカ18などが、それぞれ所定位置に配設されている。また、頭部ユニット4には、口の下顎に相当する下顎部4Aが1自由度をもって可動に取り付けられており、この下顎部4Aが動くことにより、ロボットの口の開閉動作が実現され

るようになっている。

【0018】脚部ユニット3A乃至3Dそれぞれの関節部分や、脚部ユニット3A乃至3Dそれぞれと胴体部ユニット2の連結部分、頭部ユニット4と胴体部ユニット2の連結部分、並びに尻尾部ユニット5と胴体部ユニット2の連結部分などには、図2に示すように、それぞれアクチュエータ3AA<sub>1</sub>乃至3AA<sub>K</sub>、3BA<sub>1</sub>乃至3BA<sub>K</sub>、3CA<sub>1</sub>乃至3CA<sub>K</sub>、3DA<sub>1</sub>乃至3DA<sub>K</sub>、4A<sub>1</sub>乃至4A<sub>L</sub>、5A<sub>1</sub>および5A<sub>2</sub>が配設されている。

【0019】頭部ユニット4におけるマイク15は、ユーザからの発話を含む周囲の音声（音）を集音し、得られた音声信号を、コントローラ10に送出する。CCDカメラ16は、周囲の状況を撮像し、得られた画像信号を、コントローラ10に送出する。

【0020】タッチセンサ17は、例えば、頭部ユニット4の上部に設けられており、ユーザからの「なでる」や「たたく」といった物理的な動きかけにより受けた圧力を検出し、その検出結果を圧力検出信号としてコントローラ10に送出する。

【0021】胴体部ユニット2におけるバッテリセンサ12は、バッテリ11の残量を検出し、その検出結果を、バッテリ残量検出信号としてコントローラ10に送出する。熱センサ13は、ロボット内部の熱を検出し、その検出結果を、熱検出信号としてコントローラ10に送出する。

【0022】コントローラ10は、CPU(Central Processing Unit)10Aやメモリ10B等を内蔵しており、CPU10Aにおいて、メモリ10Bに記憶された制御プログラムが実行されることにより、各種の処理を行う。

【0023】即ち、コントローラ10は、マイク15や、CCDカメラ16、タッチセンサ17、バッテリセンサ12、熱センサ13から与えられる音声信号、画像信号、圧力検出信号、バッテリ残量検出信号、熱検出信号に基づいて、周囲の状況や、ユーザからの指令、ユーザからの動きかけなどの有無を判断する。

【0024】さらに、コントローラ10は、この判断結果等に基づいて、続く行動を決定し、その決定結果に基づいて、アクチュエータ3AA<sub>1</sub>乃至3AA<sub>K</sub>、3BA<sub>1</sub>乃至3BA<sub>K</sub>、3CA<sub>1</sub>乃至3CA<sub>K</sub>、3DA<sub>1</sub>乃至3DA<sub>K</sub>、4A<sub>1</sub>乃至4A<sub>L</sub>、5A<sub>1</sub>、5A<sub>2</sub>のうちの必要なものを駆動させる。これにより、頭部ユニット4を上下左右に振らせたり、下顎部4Aを開閉させる。さらには、尻尾部ユニット5を動かせたり、各脚部ユニット3A乃至3Dを駆動して、ロボットを歩行させるなどの行動を行わせる。

【0025】また、コントローラ10は、必要に応じて、合成音を生成し、スピーカ18に供給して出力させたり、ロボットの「目」の位置に設けられた図示しない

LED (Light Emitting Diode) を点灯、消灯または点滅させる。

【0026】以上のようにして、ロボットは、周囲の状況等に基づいて自律的に行動をとるようになっている。

【0027】次に、図3は、図2のコントローラ10の機能的構成例を示している。なお、図3に示す機能的構成は、CPU10Aが、メモリ10Bに記憶された制御プログラムを実行することで実現されるようになっている。

【0028】コントローラ10は、特定の外部状態を認識するセンサ入力処理部50、センサ入力処理部50の認識結果を累積して、感情や、本能、成長の状態を表現するモデル記憶部51、センサ入力処理部50の認識結果等に基づいて、統く行動を決定する行動決定機構部52、行動決定機構部52の決定結果に基づいて、実際にロボットに行動を起こさせる姿勢遷移機構部53、各アクチュエータ3AA<sub>1</sub>乃至5A<sub>1</sub>および5A<sub>2</sub>を駆動制御する制御機構部54、並びに合成音を生成する音声合成部55から構成されている。

【0029】センサ入力処理部50は、マイク15や、CCDカメラ16、タッチセンサ17等から与えられる音声信号、画像信号、圧力検出信号等に基づいて、特定の外部状態や、ユーザからの特定の働きかけ、ユーザからの指示等を認識し、その認識結果を表す状態認識情報を、モデル記憶部51および行動決定機構部52に通知する。

【0030】即ち、センサ入力処理部50は、音声認識部50Aを有しており、音声認識部50Aは、マイク15から与えられる音声信号について音声認識を行う。そして、音声認識部50Aは、その音声認識結果としての、例えば、「歩け」、「伏せ」、「ボールを追いかけろ」等の指令その他を、状態認識情報として、モデル記憶部51および行動決定機構部52に通知する。

【0031】また、センサ入力処理部50は、画像認識部50Bを有しており、画像認識部50Bは、CCDカメラ16から与えられる画像信号を用いて、画像認識処理を行う。そして、画像認識部50Bは、その処理の結果、例えば、「赤い丸いもの」や、「地面に対して垂直なかつ所定高さ以上の平面」等を検出したときには、「ボールがある」や、「壁がある」等の画像認識結果を、状態認識情報として、モデル記憶部51および行動決定機構部52に通知する。

【0032】さらに、センサ入力処理部50は、圧力処理部50Cを有しており、圧力処理部50Cは、タッチセンサ17から与えられる圧力検出信号を処理する。そして、圧力処理部50Cは、その処理の結果、所定の閾値以上で、かつ短時間の圧力を検出したときには、「たたかれた（しかられた）」と認識し、所定の閾値未満で、かつ長時間の圧力を検出したときには、「なでられた（ほめられた）」と認識して、その認識結果を、状態

認識情報として、モデル記憶部51および行動決定機構部52に通知する。

【0033】モデル記憶部51は、ロボットの感情、本能、成長の状態を表現する感情モデル、本能モデル、成長モデルをそれぞれ記憶、管理している。

【0034】ここで、感情モデルは、例えば、「うれしさ」、「悲しさ」、「怒り」、「楽しさ」等の感情の状態（度合い）を、所定の範囲の値によってそれぞれ表し、センサ入力処理部50からの状態認識情報や時間経過等に基づいて、その値を変化させる。本能モデルは、例えば、「食欲」、「睡眠欲」、「運動欲」等の本能による欲求の状態（度合い）を、所定の範囲の値によってそれぞれ表し、センサ入力処理部50からの状態認識情報や時間経過等に基づいて、その値を変化させる。成長モデルは、例えば、「幼年期」、「青年期」、「熟年期」、「老年期」等の成長の状態（度合い）を、所定の範囲の値によってそれぞれ表し、センサ入力処理部50からの状態認識情報や時間経過等に基づいて、その値を変化させる。

【0035】モデル記憶部51は、上述のようにして感情モデル、本能モデル、成長モデルの値で表される感情、本能、成長の状態を、状態情報として、行動決定機構部52に送出する。

【0036】なお、モデル記憶部51には、センサ入力処理部50から状態認識情報が供給される他、行動決定機構部52から、ロボットの現在または過去の行動、具体的には、例えば、「長時間歩いた」などの行動の内容を示す行動情報が供給されるようになっており、モデル記憶部51は、同一の状態認識情報が与えられても、行動情報が示すロボットの行動に応じて、異なる状態情報を生成するようになっている。

【0037】即ち、例えば、ロボットが、ユーザに挨拶をし、ユーザに頭を撫でられた場合には、ユーザに挨拶をしたという行動情報と、頭を撫でられたという状態認識情報とが、モデル記憶部51に与えられ、この場合、モデル記憶部51では、「うれしさ」を表す感情モデルの値が増加される。

【0038】一方、ロボットが、何らかの仕事を実行中に頭を撫でられた場合には、仕事を実行中であるという行動情報と、頭を撫でられたという状態認識情報とが、モデル記憶部51に与えられ、この場合、モデル記憶部51では、「うれしさ」を表す感情モデルの値は変化されない。

【0039】このように、モデル記憶部51は、状態認識情報だけでなく、現在または過去のロボットの行動を示す行動情報も参照しながら、感情モデルの値を設定する。これにより、例えば、何らかのタスクを実行中に、ユーザが、いたずらするつもりで頭を撫でたときに、「うれしさ」を表す感情モデルの値を増加させるような、不自然な感情の変化が生じることを回避することが

できる。

【0040】なお、モデル記憶部51は、本能モデルおよび成長モデルについても、感情モデルにおける場合と同様に、状態認識情報および行動情報の両方に基づいて、その値を増減させるようになっている。また、モデル記憶部51は、感情モデル、本能モデル、成長モデルそれぞれの値を、他のモデルの値にも基づいて増減させるようになっている。

【0041】行動決定機構部52は、センサ入力処理部50からの状態認識情報や、モデル記憶部51からの状態情報、時間経過等に基づいて、次の行動を決定し、決定された行動の内容を、行動指令情報として、姿勢遷移機構部53に送出する。

【0042】即ち、行動決定機構部52は、ロボットがとり得る行動をステート（状態）(state)に対応させた有限オートマトンを、ロボットの行動を規定する行動モデルとして管理しており、この行動モデルとしての有限オートマトンにおけるステートを、センサ入力処理部50からの状態認識情報や、モデル記憶部51における感情モデル、本能モデル、または成長モデルの値、時間経過等に基づいて遷移させ、遷移後のステートに対応する行動を、次にとるべき行動として決定する。

【0043】ここで、行動決定機構部52は、所定のトリガ(trigger)があったことを検出すると、ステートを遷移させる。即ち、行動決定機構部52は、例えば、現在のステートに対応する行動を実行している時間が所定時間に達したときや、特定の状態認識情報を受信したとき、モデル記憶部51から供給される状態情報が示す感情や、本能、成長の状態の値が所定の閾値以下または以上になったとき等に、ステートを遷移させる。

【0044】なお、行動決定機構部52は、上述したように、センサ入力処理部50からの状態認識情報だけでなく、モデル記憶部51における感情モデルや、本能モデル、成長モデルの値等にも基づいて、行動モデルにおけるステートを遷移させることから、同一の状態認識情報が入力されても、感情モデルや、本能モデル、成長モデルの値（状態情報）によっては、ステートの遷移先は異なるものとなる。

【0045】その結果、行動決定機構部52は、例えば、状態情報が、「怒っていない」こと、および「お腹がすいていない」ことを表している場合において、状態認識情報が、「目の前に手のひらが差し出された」ことを表しているときには、目の前に手のひらが差し出されたことに応じて、「お手」という行動をとらせる行動指令情報を生成し、これを、姿勢遷移機構部53に送出する。

【0046】また、行動決定機構部52は、例えば、状態情報が、「怒っていない」こと、および「お腹がすいている」ことを表している場合において、状態認識情報が、「目の前に手のひらが差し出された」ことを表して

いるときには、目の前に手のひらが差し出されたことに応じて、「手のひらをぺろぺろなめる」ような行動を行わせるための行動指令情報を生成し、これを、姿勢遷移機構部53に送出する。

【0047】また、行動決定機構部52は、例えば、状態情報が、「怒っている」ことを表している場合において、状態認識情報が、「目の前に手のひらが差し出された」ことを表しているときには、状態情報が、「お腹がすいている」ことを表していても、また、「お腹がすいていない」ことを表していても、「ぶいと横を向く」ような行動を行わせるための行動指令情報を生成し、これを、姿勢遷移機構部53に送出する。

【0048】なお、行動決定機構部52には、モデル記憶部51から供給される状態情報が示す感情や、本能、成長の状態に基づいて、遷移先のステートに対応する行動のパラメータとしての、例えば、歩行の速度や、手足を動かす際の動きの大きさおよび速度などを決定させることができ、この場合、それらのパラメータを含む行動指令情報が、姿勢遷移機構部53に送出される。

【0049】また、行動決定機構部52では、上述したように、ロボットの頭部や手足等を動作させる行動指令情報の他、ロボットに発話を行わせる行動指令情報も生成される。ロボットに発話を行わせる行動指令情報は、音声合成部55に供給されるようになっており、音声合成部55に供給される行動指令情報には、音声合成部55に生成させる合成音に対応するテキスト等が含まれる。そして、音声合成部55は、行動決定部52から行動指令情報を受信すると、その行動指令情報に含まれるテキストに基づき、合成音を生成し、スピーカ18に供給して出力させる。これにより、スピーカ18からは、例えば、ロボットの鳴き声、さらには、「お腹がすいた」等のユーザへの各種の要求、「何？」等のユーザの呼びかけに対する応答その他の音声出力が行われる。ここで、音声合成部55には、モデル記憶部51から状態情報も供給されるようになっており、音声合成部55は、この状態情報が示す感情の状態に基づいて韻律を制御した合成音を生成することが可能となっている。なお、音声合成部55では、感情の他、本能や成長の状態に基づいて韻律を制御した合成音を生成することも可能である。

【0050】姿勢遷移機構部53は、行動決定機構部52から供給される行動指令情報に基づいて、ロボットの姿勢を、現在の姿勢から次の姿勢に遷移させるための姿勢遷移情報を生成し、これを制御機構部54に送出する。

【0051】ここで、現在の姿勢から次に遷移可能な姿勢は、例えば、胴体や手や足の形状、重さ、各部の結合状態のようなロボットの物理的形状と、関節が曲がる方向や角度のようなアクチュエータ3AA<sub>1</sub>乃至5A<sub>1</sub>および5A<sub>2</sub>の機構とによって決定される。

【0052】また、次の姿勢としては、現在の姿勢から直接遷移可能な姿勢と、直接には遷移できない姿勢がある。例えば、4本足のロボットは、手足を大きく投げ出して寝転んでいる状態から、伏せた状態へ直接遷移することはできるが、立った状態へ直接遷移することはできず、一旦、手足を胴体近くに引き寄せて伏せた姿勢になり、それから立ち上がるという2段階の動作が必要である。また、安全に実行できない姿勢も存在する。例えば、4本足のロボットは、その4本足で立っている姿勢から、両前足を挙げてバンザイをしようとする、簡単に転倒してしまう。

【0053】このため、姿勢遷移機構部53は、直接遷移可能な姿勢をあらかじめ登録しておき、行動決定機構部52から供給される行動指令情報が、直接遷移可能な姿勢を示す場合には、その行動指令情報を、そのまま姿勢遷移情報として、制御機構部54に送出する。一方、行動指令情報が、直接遷移不可能な姿勢を示す場合には、姿勢遷移機構部53は、遷移可能な他の姿勢に一旦遷移した後に、目的の姿勢まで遷移させるような姿勢遷移情報を生成し、制御機構部54に送出する。これによりロボットが、遷移不可能な姿勢を無理に実行しようとする事態や、転倒するような事態を回避することができるようになっている。

【0054】制御機構部54は、姿勢遷移機構部53からの姿勢遷移情報にしたがって、アクチュエータ3AA<sub>1</sub>乃至5A<sub>1</sub>および5A<sub>2</sub>を駆動するための制御信号を生成し、これを、アクチュエータ3AA<sub>1</sub>乃至5A<sub>1</sub>および5A<sub>2</sub>に送出する。これにより、アクチュエータ3AA<sub>1</sub>乃至5A<sub>1</sub>および5A<sub>2</sub>は、制御信号にしたがって駆動し、ロボットは、自律的に行動を起こす。

【0055】次に、図4は、図3の音声認識部50Aの構成例を示している。

【0056】マイク15からの音声信号は、AD(Analog Digital)変換部21に供給される。AD変換部21では、マイク15からのアナログ信号である音声信号がサンプリング、量子化され、ディジタル信号である音声データにA/D変換される。この音声データは、特徴抽出部22および音声区間検出部27に供給される。

【0057】特徴抽出部22は、そこに入力される音声データについて、適当なフレームごとに、例えば、MFCC(Mel Frequency Cepstrum Coefficient)分析を行い、その分析の結果得られるMFCCを、特徴パラメータ(特徴ベクトル)として、マッチング部23に出力する。なお、特徴抽出部22では、その他、例えば、線形予測係数、ケプストラム係数、線スペクトル対、所定の周波数帯域ごとのパワー(フィルタバンクの出力)等を、特徴パラメータとして抽出することが可能である。

【0058】マッチング部23は、特徴抽出部22からの特徴パラメータを用いて、音響モデル記憶部24、辞書記憶部25、および文法記憶部26を必要に応じて参

照しながら、マイク15に入力された音声(入力音声)を、例えば、連続分布HMM(Hidden Markov Model)法に基づいて音声認識する。

【0059】即ち、音響モデル記憶部24は、音声認識する音声の言語における個々の音素や音節などの音響的な特徴を表す音響モデルを記憶している。ここでは、連続分布HMM法に基づいて音声認識を行うので、音響モデルとしては、HMM(Hidden Markov Model)が用いられる。辞書記憶部25は、認識対象の各単語について、その発音に関する情報(音韻情報)が記述された単語辞書を記憶している。文法記憶部26は、辞書記憶部25の単語辞書に登録されている各単語が、どのように連鎖する(つながる)かを記述した文法规則を記憶している。ここで、文法规則としては、例えば、文脈自由文法(CFG)や、統計的な単語連鎖確率(N-gram)などに基づく規則を用いることができる。

【0060】マッチング部23は、辞書記憶部25の単語辞書を参照することにより、音響モデル記憶部24に記憶されている音響モデルを接続することで、単語の音響モデル(単語モデル)を構成する。さらに、マッチング部23は、幾つかの単語モデルを、文法記憶部26に記憶された文法规則を参照することにより接続し、そのようにして接続された単語モデルを用いて、特徴パラメータに基づき、連続分布HMM法によって、マイク15に入力された音声を認識する。即ち、マッチング部23は、特徴抽出部22が出力する時系列の特徴パラメータが観測されるスコア(尤度)が最も高い単語モデルの系列を検出し、その単語モデルの系列に対応する単語列の音韻情報(読み)を、音声の認識結果として出力する。

【0061】より具体的には、マッチング部23は、接続された単語モデルに対応する単語列について、各特徴パラメータの出現確率を累積し、その累積値をスコアとして、そのスコアを最も高くする単語列の音韻情報を、音声認識結果として出力する。

【0062】以上のようにして出力される、マイク15に入力された音声の認識結果は、状態認識情報として、モデル記憶部51および行動決定機構部52に出力される。

【0063】なお、音声区間検出部27は、AD変換部21からの音声データについて、特徴抽出部22がMFCC分析を行うのと同様のフレームごとに、例えば、パワーを算出している。さらに、音声区間検出部27は、各フレームのパワーを、所定の閾値と比較し、その閾値以上のパワーを有するフレームで構成される区間を、ユーザの音声が入力されている音声区間として検出する。そして、音声区間検出部27は、検出した音声区間を、特徴抽出部22とマッチング部23に供給しており、特徴抽出部22とマッチング部23は、音声区間のみを対象に処理を行う。

【0064】次に、図5は、図3の音声合成部55の構

成例を示している。

【0065】テキスト解析部31には、行動決定機構部52が outputする、音声合成の対象とするテキストを含む行動指令情報が供給されるようになっており、テキスト解析部31は、辞書記憶部34や生成用文法記憶部35を参照しながら、その行動指令情報に含まれるテキストを解析する。

【0066】即ち、辞書記憶部34には、各単語の品詞情報や、読み、アクセント等の情報が記述された単語辞書が記憶されており、また、生成用文法記憶部35には、辞書記憶部34の単語辞書に記述された単語について、単語連鎖に関する制約等の生成用文法規則が記憶されている。そして、テキスト解析部31は、この単語辞書および生成用文法規則に基づいて、そこに入力されるテキストの形態素解析や構文解析等のテキスト解析を行い、後段の規則合成部32で行われる規則音声合成に必要な情報を抽出する。ここで、規則音声合成に必要な情報としては、例えば、ポーズの位置や、アクセント、イントネーション、パワー等を制御するための情報、各単語の発音を表す音韻情報などがある。

【0067】テキスト解析部31で得られた情報は、規則合成部32に供給され、規則合成部32は、音素片記憶部36を参照しながら、テキスト解析部31に入力されたテキストに対応する合成音の音声データ（デジタルデータ）を生成する。

【0068】即ち、音素片記憶部36には、例えば、C V (Consonant, Vowel) や、VCV, CVC、あるいは1ピッチ等の形で音素片データが記憶されており、規則合成部32は、テキスト解析部31からの情報に基づいて、必要な音素片データを接続し、さらに、音素片データの波形を加工することによって、ポーズ、アクセント、イントネーション等を適切に付加し、これにより、テキスト解析部31に入力されたテキストに対応する合成音の音声データ（合成音データ）を生成する。

【0069】また、規則合成部32には、モデル記憶部51から状態情報が供給されるようになっており、規則合成部32は、その状態情報のうちの、例えば、感情モデルの値に基づいて韻律を制御した合成音データを生成する。

【0070】以上のようにして生成された合成音データは、スピーカ18に供給され、これにより、スピーカ18からは、テキスト解析部31に入力されたテキストに対応する合成音が出力される。

【0071】なお、図3の行動決定機構部52では、上述したように、行動モデルに基づいて、次の行動が決定されるが、合成音として出力するテキストの内容は、ロボットの行動と対応付けておくことが可能である。

【0072】即ち、例えば、ロボットが、座った状態から、立った状態になる行動には、テキスト「よっこいしょ」などを対応付けておくことが可能である。この場

合、ロボットが、座っている姿勢から、立つ姿勢に移行するときに、その姿勢の移行に同期して、合成音「よっこいしょ」を出力することが可能となる。

【0073】次に、図6は、図5の規則音声合成部32の構成例を示している。

【0074】韻律生成部41には、テキスト解析部31（図5）によるテキスト解析結果が供給され、韻律生成部41は、そのテキスト解析結果に基づいて、合成音の韻律を制御する韻律情報を生成する。ここで、韻律制御部41では、合成音を構成する各音韻の継続時間長、合成音のピッチ周期の時間変化パターンを表す周期パターン信号、合成音のパワーの時間変化パターンを表すパワーパターン信号等が、韻律情報として生成される。

【0075】韻律生成部41で生成された継続時間長は、波形生成部44に供給される。また、周期パターン信号は、演算器42に供給され、パワーパターン信号は、演算器43に供給される。

【0076】演算器42は、韻律生成部41から供給される周期パターン信号と、周期パターン用周期信号生成部46から供給される周期パターン用周期信号とを加算することにより、周期パターン信号を補正し、その補正後の周期パターン信号（以下、適宜、補正周期パターン信号という）を、波形生成部44に供給する。演算器43は、韻律生成部41から供給されるパワーパターン信号と、パワーパターン用周期信号生成部47から供給されるパワーパターン用周期信号とを加算することにより、パワーパターン信号を補正し、その補正後のパワーパターン信号（以下、適宜、補正パワーパターン信号という）を、波形生成部44に供給する。

【0077】波形生成部44には、上述したように、韻律情報としての継続時間長、補正周期パターン信号、および補正パワーパターン信号が供給される他、テキスト解析部31（図5）によるテキスト解析結果が供給される。波形生成部41は、テキスト解析結果に含まれる音韻情報にしたがって、必要な音素片データを、音素片記憶部36（図5）から読み出して接続し、さらに、韻律情報としての継続時間長、補正周期パターン信号、および補正パワーパターン信号にしたがって、接続した音素片データの波形を調整（成形）し、これにより、最終的な合成音データを生成して出力する。

【0078】制御信号生成部45には、モデル記憶部51から状態情報が供給されるようになっており、制御信号生成部45は、その状態情報のうちの感情モデルに基づいて、周期パターン用周期信号の周期と振幅を制御する制御信号（以下、適宜、周期パターン用制御信号という）、およびパワーパターン用周期信号の周期と振幅を制御する制御信号（以下、適宜、パワーパターン用制御信号という）を生成する。

【0079】即ち、制御信号生成部45は、例えば、感情モデルとしての「うれしさ」、「悲しさ」、「怒

り」、「楽しさ」等の感情の状態を表す値（以下、適宜、感情モデル値という）に、周期パターン用周期信号の周期と振幅を対応付けた変換テーブルを記憶しており、その変換テーブルにおいて、モデル記憶部51からの状態情報における感情モデル値に対応付けられている周期と振幅を表す周期パターン用制御信号を出力する。さらに、制御信号生成部45は、パワーパターン用周期信号についても同様の変換テーブルを記憶しており、その変換テーブルを参照することにより、モデル記憶部51からの状態情報における感情モデル値に対応付けられている周期と振幅を表すパワーパターン用制御信号を出力する。

【0080】なお、制御信号生成部45が記憶している変換テーブルにおいては、ペットロボットの感情の状態を表すような韻律の合成音が得られるように、感情モデル値と、周期および振幅とが対応付けられている。即ち、例えば、「怒り」の感情が大のときには、パワーパターン用周期信号の振幅が大となるように、また、例えば、「悲しさ」の感情が大のときには、周期パターン用周期信号の周期が大となるように、変換テーブルは構成されている。ここで、感情モデル値と、周期および振幅とを、どのように対応付けるかは、例えば、シミュレーションを行い、そのシミュレーション結果に基づいて決定することができる。

【0081】また、上述の場合には、「うれしさ」、「悲しさ」、「怒り」、「楽しさ」等のすべての感情モデル値を考慮した変換テーブルを用いるようにしたが、その他、例えば、次のような簡略化した変換テーブルを用いることも可能である。

【0082】即ち、感情の状態を、例えば、「定常」、「悲しさ」、「怒り」、「楽しさ」等のいずれかのみに分類し、各感情に、ユニークな番号としての感情番号を付しておく。具体的には、例えば、「定常」、「悲しさ」、「怒り」、「楽しさ」に、それぞれ0, 1, 2, 3等の感情番号を、それぞれ付しておく。そして、このような感情番号と、周期および振幅とを対応付けた変換テーブルを作成する。なお、このような変換テーブルを用いる場合には、感情モデル値から、感情の状態を、「うれしさ」、「悲しさ」、「怒り」、「楽しさ」のいずれかに分類する必要があるが、これは、次のようにして行うことが可能である。即ち、例えば、複数の感情モデル値のうち、最も大きい感情モデル値と、2番目に大きい感情モデル値との差が、所定の閾値以上の場合は、最も大きい感情モデル値に対応する感情の状態に分類し、そうでない場合は、「定常」の状態に分類すればよい。

【0083】周期パターン用周期信号生成部46は、制御信号生成部45が出力する周期パターン用制御信号が表す周期と振幅の周期的な信号を生成し、これを、周期パターン用周期信号として、演算器42に出力する。パ

ワーパターン用周期信号生成部47は、制御信号生成部45が出力するパワーパターン用制御信号が表す周期と振幅の周期的な信号を生成し、これを、パワーパターン用周期信号として、演算器43に出力する。

【0084】なお、周期パターン用周期信号と、パワーパターン用周期信号としては、例えば、正弦波や三角波等を用いることができる。

【0085】次に、図7のフローチャートを参照して、図6の規則合成部32の処理について説明する。

【0086】図5のテキスト解析部31（図5）が出力するテキスト解析結果は、韻律生成部31と波形生成部44に供給される。また、図5のモデル記憶部51が出力する状態情報は、制御信号生成部45に供給される。

【0087】韻律生成部41は、テキスト解析結果を受信すると、ステップS1において、韻律情報として、テキスト解析結果に含まれる音韻情報が表す各音韻の継続時間長、周期パターン信号、パワーパターン信号を生成する。継続時間長は、波形生成部44に供給され、周期パターン信号とパワーパターン信号は、それぞれ、演算器42と43に供給される。

【0088】その後、ステップS2に進み、制御信号生成部45は、感情反映モードかどうかを判定する。即ち、本実施の形態では、感情を反映した合成音を出力する感情反映モードと、感情を反映しない合成音を出力する非感情反映モードのうちのいずれかを設定することができるようになっており、ステップS2では、ロボットのモードが感情反映モードとなっているかどうかが判定される。

【0089】ここで、ロボットには、感情反映モードと非感情反映モードを設けずに、常に、感情を反映した合成音を出力させることも可能である。

【0090】ステップS2において、感情反映モードでないと判定された場合、ステップS3乃至S5をスキップして、ステップS6に進み、波形生成部44は、合成音を生成する。

【0091】即ち、感情反映モードでない場合、制御信号生成部44は、特に処理を行わず、従って、周期パターン用周期信号生成部46とパワーパターン用周期信号生成部47は、それぞれ、周期パターン用周期信号とパワーパターン用周期信号を出力しない。

【0092】その結果、演算器42と43では、それぞれ、韻律生成部41が出力する周期パターン信号とパワーパターン信号が、特に処理されず、そのまま、波形生成部44に出力される。

【0093】従って、波形生成部44は、韻律生成部41で生成された継続時間長、周期パターン信号、およびパワーパターン信号によって与えられる韻律を有する合成音データが生成される。

【0094】即ち、波形生成部44は、テキスト解析部31（図5）からのテキスト解析結果に含まれる音韻情

報にしたがって、必要な音素片データを、音素片記憶部36(図5)から読み出して接続し、さらに、韻律情報としての継続時間長、周期パターン信号、およびパワーパターン信号にしたがって、接続した音素片データの波形を調整し、これにより、最終的な合成音データを生成して出力する。

【0095】一方、ステップS2において、感情反映モードであると判定された場合、ステップS3に進み、制御信号生成部45は、モデル記憶部51からの状態情報のうちの感情モデル値に基づいて、周期パターン用周期信号の周期と振幅を制御する周期パターン用制御信号、およびパワーパターン用周期信号の周期と振幅を制御するパワーパターン用制御信号を生成する。この周期パターン用制御信号は、周期パターン用周期信号生成部46に供給され、パワーパターン用制御信号は、パワーパターン用周期信号生成部47に供給される。

【0096】その後、ステップS4に進み、周期パターン用周期信号生成部46が、制御信号生成部45からの周期パターン用制御信号が表す周期と振幅の周期パターン用周期信号を生成するとともに、パワーパターン用周期信号生成部47が、同じく、制御信号生成部45からのワーパターン用制御信号が表す周期と振幅のパワーパターン用周期信号を生成する。周期パターン用周期信号は、演算器42に供給され、パワーパターン用周期信号は、演算器43に供給される。

【0097】そして、ステップS5に進み、演算器42が、韻律生成部41からの周期パターン信号と、周期パターン用周期信号生成部46からの周期パターン用周期信号とを加算することにより、周期パターン信号を補正し、その補正後の周期パターン信号である補正周期パターン信号を、波形生成部44に供給する。

【0098】さらに、ステップS5では、演算器43が、韻律生成部41からのパワーパターン信号と、パワーパターン用周期信号生成部47からのパワーパターン用周期信号とを加算することにより、パワーパターン信号を補正し、その補正後のパワーパターン信号である補正パワーパターン信号を、波形生成部44に供給する。

【0099】即ち、ステップS5では、例えば、図8(A)に示すような周期パターン信号と、図8(B)に示すような周期パターン用周期信号とを加算することにより、図8(C)に示すような補正周期パターン信号が生成される。補正パワーパターン信号も同様にして生成される。

【0100】波形生成部41は、韻律生成部41から継続時間長を受信するとともに、演算器42と43からそれぞれ補正周期パターン信号と補正パワーパターン信号を受信すると、ステップS6において、合成音を生成して、処理を終了する。

【0101】即ち、波形生成部44は、テキスト解析部31(図5)からのテキスト解析結果に含まれる音韻情

報にしたがって、必要な音素片データを、音素片記憶部36(図5)から読み出して接続し、さらに、韻律情報としての継続時間長、補正周期パターン信号、および補正パワーパターン信号にしたがって、接続した音素片データの波形を調整し、これにより、最終的な合成音データを生成して出力する。

【0102】以上のように、感情モデル値に基づき、周期パターン(またはパワーパターン)の韻律情報を周期的に変化させる周期パターン用周期信号(またはパワーパターン用周期信号)を生成し、その周期パターン用周期信号(またはパワーパターン用周期信号)と、周期パターン信号(またはパワーパターン信号)とを加算することにより、周期パターン信号(またはパワーパターン信号)を補正し、その結果得られる補正周期パターン信号(または補正パワーパターン信号)にしたがった韻律の合成音を生成するようにしたので、感情に応じて韻律が被打つように変化する、感情豊かな合成音を、容易に得ることができる。

【0103】以上、本発明を、エンターテイメント用のロボット(疑似ペットとしてのロボット)に適用した場合について説明したが、本発明は、このような現実世界のロボットだけでなく、例えば、液晶ディスプレイ等の表示装置に表示される仮想的なロボット(キャラクタ)にも適用可能である。さらに、本発明は、ロボットの他、音声合成装置を搭載した、例えば対話システムその他にも適用可能である。

【0104】なお、本実施の形態においては、上述した一連の処理を、CPU10Aにプログラムを実行させることにより行なうようにしたが、一連の処理は、それ専用のハードウェアによって行なうことも可能である。

【0105】ここで、プログラムは、あらかじめメモリ10B(図2)に記憶させておく他、フロッピー(登録商標)ディスク、CD-ROM(Compact Disc Read Only Memory), MD(Magnetooptical)ディスク、DVD(Digital Versatile Disc)、磁気ディスク、半導体メモリなどのリムーバブル記録媒体に、一時的あるいは永続的に格納(記録)しておくことができる。そして、このようなりムーバブル記録媒体を、いわゆるパッケージソフトウェアとして提供し、ロボット(メモリ10B)にインストールするようにすることができる。

【0106】また、プログラムは、ダウンロードサイトから、デジタル衛星放送用の人工衛星を介して、無線で転送したり、LAN(Local Area Network)、インターネットといったネットワークを介して、有線で転送し、メモリ10Bにインストールすることができる。

【0107】この場合、プログラムがバージョンアップされたとき等に、そのバージョンアップされたプログラムを、メモリ10Bに、容易にインストールすることができる。

【0108】なお、本明細書において、CPU10Aに

各種の処理を行わせるためのプログラムを記述する処理ステップは、必ずしもフローチャートとして記載された順序に沿って時系列に処理する必要はなく、並列的あるいは個別に実行される処理（例えば、並列処理あるいはオブジェクトによる処理）も含むものである。

【0109】また、プログラムは、1のCPUにより処理されるものであっても良いし、複数のCPUによって分散処理されるものであっても良い。

【0110】次に、図5の音声合成装置5は、専用のハードウェアにより実現することもできるし、ソフトウェアにより実現することもできる。音声合成装置5をソフトウェアによって実現する場合には、そのソフトウェアを構成するプログラムが、汎用のコンピュータ等にインストールされる。

【0111】そこで、図9は、音声合成装置5を実現するためのプログラムがインストールされるコンピュータの一実施の形態の構成例を示している。

【0112】プログラムは、コンピュータに内蔵されている記録媒体としてのハードディスク105やROM103に予め記録しておくことができる。

【0113】あるいはまた、プログラムは、フロッピーディスク、CD-ROM、MDディスク、DVD、磁気ディスク、半導体メモリなどのリムーバブル記録媒体111に、一時的あるいは永続的に格納（記録）しておくことができる。このようなリムーバブル記録媒体111は、いわゆるパッケージソフトウェアとして提供することができる。

【0114】なお、プログラムは、上述したようなリムーバブル記録媒体111からコンピュータにインストールする他、ダウンロードサイトから、ディジタル衛星放送用の人工衛星を介して、コンピュータに無線で転送したり、LAN、インターネットといったネットワークを介して、コンピュータに有線で転送し、コンピュータでは、そのようにして転送されてくるプログラムを、通信部108で受信し、内蔵するハードディスク105にインストールすることができる。

【0115】コンピュータは、CPU102を内蔵している。CPU102には、バス101を介して、入出力インターフェース110が接続されており、CPU102は、入出力インターフェース110を介して、ユーザによって、キーボードや、マウス、マイク等で構成される入力部107が操作等されることにより指令が入力されると、それにしたがって、ROM103に格納されているプログラムを実行する。あるいは、また、CPU102は、ハードディスク105に格納されているプログラム、衛星若しくはネットワークから転送され、通信部108で受信されてハードディスク105にインストールされたプログラム、またはドライブ109に装着されたリムーバブル記録媒体111から読み出されてハードディスク105にインストールされたプログラムを、RAM(Random Access

Memory)104にロードして実行する。これにより、CPU102は、上述したフローチャートにしたがった処理、あるいは上述したプログラムの構成により行われる処理を行う。そして、CPU102は、その処理結果を、必要に応じて、例えば、入出力インターフェース110を介して、LCD(Liquid Crystal Display)やスピーカ等で構成される出力部106から出力、あるいは、通信部108から送信、さらには、ハードディスク105に記録等させる。

【0116】なお、本実施の形態では、周期パターン信号とパワーパターン信号を補正するようにしたが、それ以外に、継続時間長その他の韻律情報を補正することも可能である。

【0117】また、本実施の形態では、感情の状態に基づいて、周期パターン用制御信号や、パワーパターン用制御信号を生成するようにしたが、感情以外の、例えば、ロボットの本能や成長の状態、さらには、他のパラメータに基づいて、周期パターン用制御信号や、パワーパターン用制御信号を生成することも可能である。

【0118】さらに、本実施の形態においては、周期パターン用周期信号やパワーパターン用周期信号の周期については、特に言及しなかったが、周期パターン用周期信号やパワーパターン用周期信号の周期は、あまり短くすると、合成音に、ゆらぎとなって現れることから、あまり短くしそうないようにするのが望ましい。即ち、周期パターン用周期信号やパワーパターン用周期信号の周期は、ゆらぎとならないような、例えば、音節の継続時間以上とするのが望ましい。

【0119】また、本実施の形態では、テキストから合成音を生成するようにしたが、その他、発音記号等から合成音を生成することも可能である。

【0120】

【発明の効果】本発明の音声合成装置および音声合成方法、並びにプログラムによれば、韻律情報を周期的に変化させる周期信号が生成され、周期信号と韻律情報を加算することにより、韻律情報が補正される。そして、その結果得られる補正韻律情報を用いて音声合成が行われる。従って、合成音の韻律を容易に制御することが可能となり、これにより、感情等の状態に応じた合成音を、容易に生成することが可能となる。

【図面の簡単な説明】

【図1】本発明を適用したロボットの一実施の形態の外観構成例を示す斜視図である。

【図2】ロボットの内部構成例を示すブロック図である。

【図3】コントローラ10の機能的構成例を示すブロック図である。

【図4】音声認識部50Aの構成例を示すブロック図である。

【図5】音声合成部55の構成例を示すブロック図である。

る。

【図6】規則合成部32の構成例を示すブロック図である。

【図7】規則合成部32の処理を説明するフローチャートである。

【図8】図7のステップS5の処理を説明するための図である。

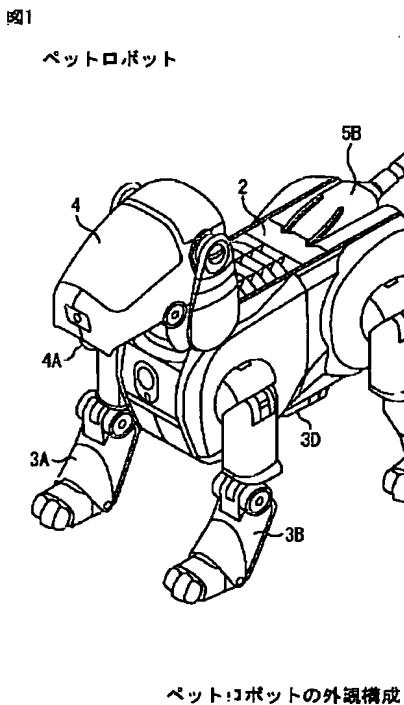
【図9】本発明を適用したコンピュータの一実施の形態の構成例を示すブロック図である。

#### 【符号の説明】

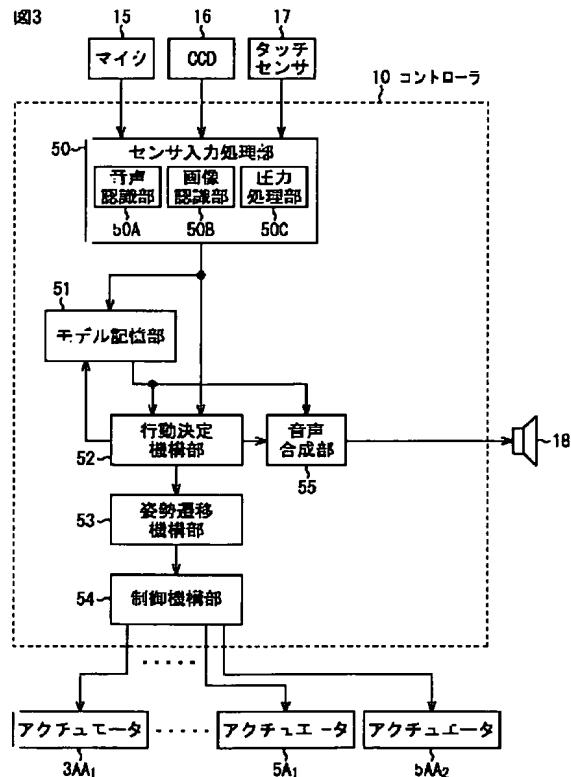
1 頭部ユニット, 4A 下顎部, 10 コントローラ, 10A CPU, 10B メモリ, 15 マイク, 16 CCDカメラ, 17 タッチセンサ, 18 スピーカ, 21 AD変換部, 22 特徴抽出部, 23 マッチング部, 24 音響モデル記憶部, 25 辞書記憶部, 26 文法記憶部,

27 音声区間検出部, 31 テキスト解析部, 32 規則合成部, 34 辞書記憶部, 35 生成用文法記憶部, 36 音素片記憶部, 41 韻律生成部, 42, 43 演算器, 44 波形生成部, 45 制御信号生成部, 46 周期パターン用周期信号生成部, 47 パワーパターン用周期信号生成部, 50 センサ入力処理部, 50A 音声認識部, 50B 画像認識部, 50C 圧力処理部, 51 モデル記憶部, 52 行動決定機構部, 53 姿勢遷移機構部, 54 制御機構部, 55 音声合成部, 101 バス, 102 CPU, 103 ROM, 104 RAM, 105 ハードディスク, 106 出力部, 107 入力部, 108 通信部, 109 ドライバ, 110 入出力インターフェース, 111 リムーバブル記録媒体

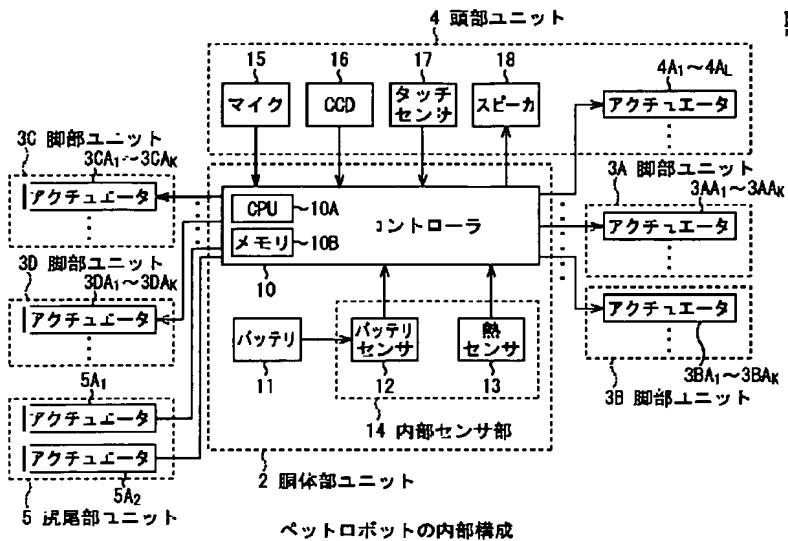
【図1】



【図3】

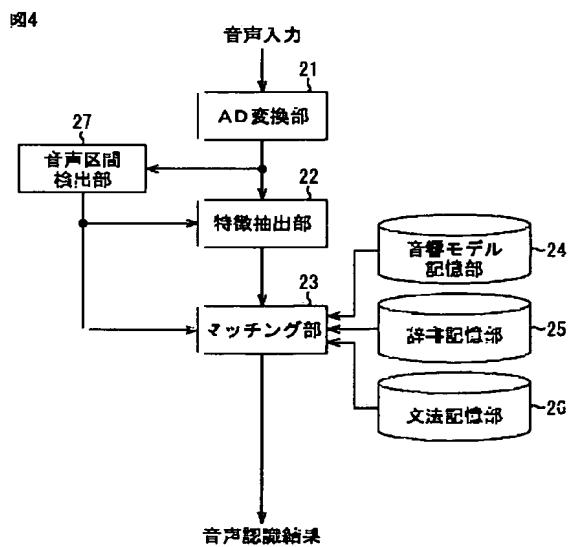


【図2】

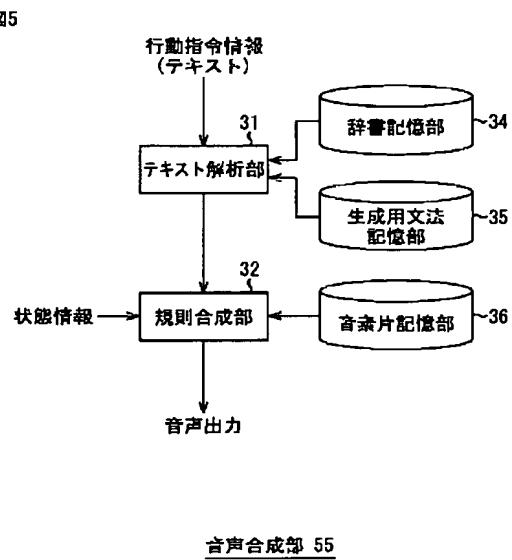


四

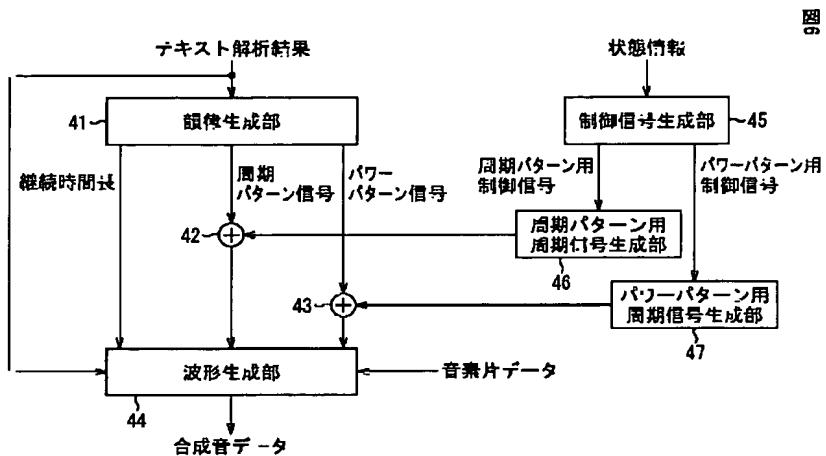
【図4】



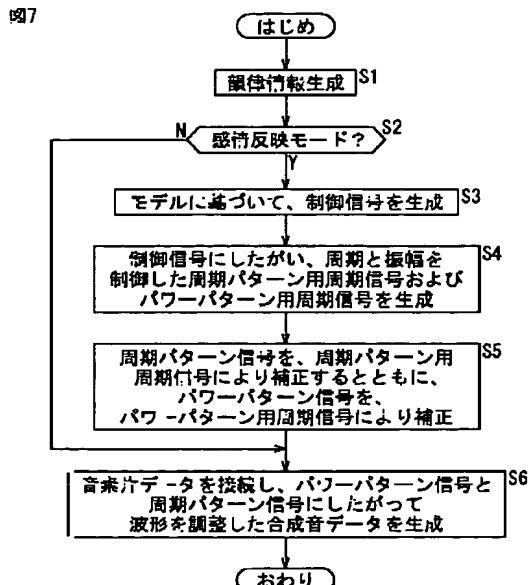
【図5】

音声認識部 50A音声合成部 55

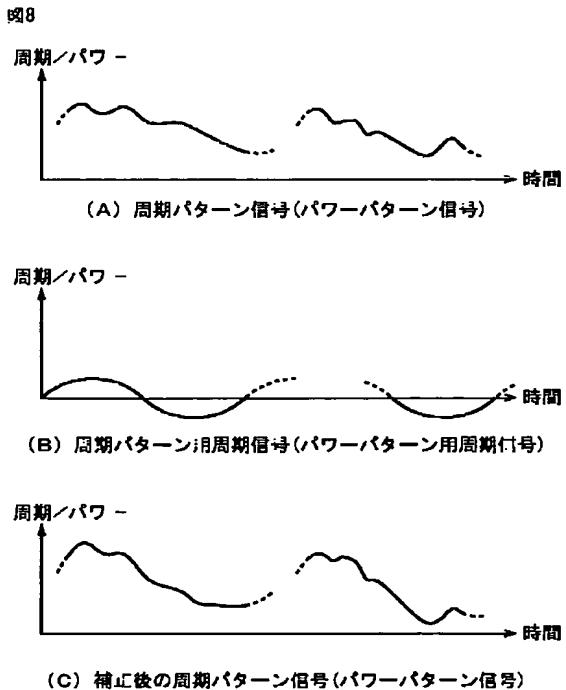
【図6】

規則合成部 32

【図7】

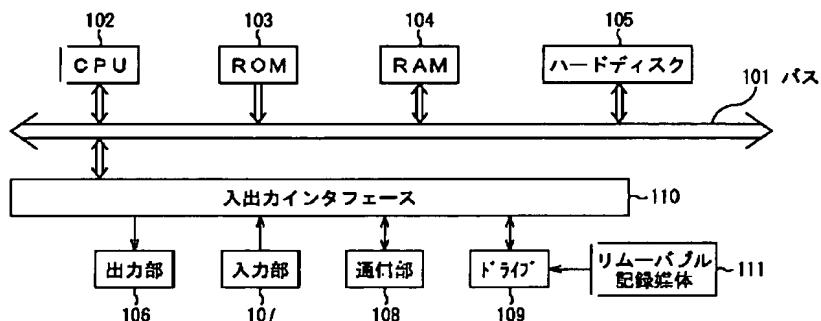


【図8】



【図9】

図



コンピュータ

## フロントページの続き

(72)発明者 山田 敬一  
東京都品川区北品川6丁目7番35号 ソニ  
一株式会社内

(72)発明者 小林 賢一郎  
東京都品川区北品川6丁目7番35号 ソニ  
一株式会社内

(72)発明者 新田 朋晃  
東京都品川区北品川6丁目7番35号 ソニ  
一株式会社内

(72)発明者 小林 恵理香  
東京都品川区北品川6丁目7番35号 ソニ  
一株式会社内

Fターム(参考) 2C150 BA11 CA01 CA02 CA04 DA04  
DA05 DA24 DA25 DA26 DA27  
DA28 DF03 DF04 DF06 DF33  
ED42 ED52 EF07 EF13 EF16  
EF23 EF29 EF34 EF36  
3C007 AS36 CS08 MT14 WA04 WA14  
WB19 WB22 WB27 WC30  
5D045 AA09 AB11